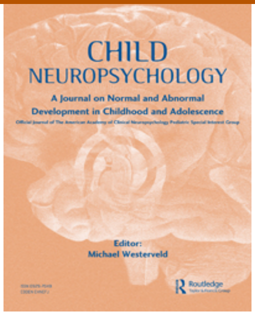

Child Neuropsychology

A Journal on Normal and Abnormal Development in Childhood and Adolescence

ISSN: 0929-7049 (Print) 1744-4136 (Online) Journal homepage: http://www.tandfonline.com/loi/ncny20

Inter-individual differences in how presentation modality affects verbal learning performance in children aged 5 to 16

Celeste Meijs, Petra P. M. Hurks, Renske Wassenberg, Frans J. M. Feron & Jelle Jolles

To cite this article: Celeste Meijs, Petra P. M. Hurks, Renske Wassenberg, Frans J. M. Feron & Jelle Jolles (2016) Inter-individual differences in how presentation modality affects verbal learning performance in children aged 5 to 16, Child Neuropsychology, 22:7, 818-836, DOI: 10.1080/09297049.2015.1044891

To link to this article: http://dx.doi.org/10.1080/09297049.2015.1044891

Published online: 22 Jun 2015.

Child Neuropsychology, 2016
Vol. 22, No. 7, 818–836, http://dx.doi.org/10.1080/09297049.2015.1044891

Inter-individual differences in how presentation modality affects verbal learning performance in children aged 5 to 16

Celeste Meijs[1,4], Petra P. M. Hurks[2], Renske Wassenberg[3], Frans J. M. Feron[4], and Jelle Jolles[5]

[1]Welten Institute, Research Centre for Learning, Teaching and Technology, Open University, The Netherlands
[2]Department of Neuropsychology & Psychopharmacology, Faculty of Psychology & Neuroscience, Maastricht University, Maastricht, The Netherlands
[3]Department of Psychiatry & Neuropsychology, Academic Hospital Maastricht, Maastricht, The Netherlands
[4]Faculty of Health, Medicine &Life Sciences, Maastricht University, Maastricht, The Netherlands
[5]LEARN! Research Institute, Faculty of Psychology & Education, VU University Amsterdam, Amsterdam, The Netherlands

This study examines inter-individual differences in how presentation modality affects verbal learning performance. Children aged 5 to 16 performed a verbal learning test within one of three presentation modalities: pictorial, auditory, or textual. The results indicated that a beneficial effect of pictures exists over auditory and textual presentation modalities and that this effect increases with age. However, this effect is only found if the information to be learned is presented once (or at most twice) and only in children above the age of 7. The results may be explained in terms of single or dual coding of information in which the phonological loop is involved. Development of the (sub)vocal rehearsal system in the phonological loop is believed to be a gradual process that begins developing around the age of 7. The developmental trajectories are similar for boys and girls. Additionally, auditory information and textual information both seemed to be processed in a similar manner, namely without labeling or recoding, leading to single coding. In contrast, pictures are assumed to be processed by the dual coding of both the visual information and a (verbal) labeling of the pictures.

***Keywords*:** Verbal Learning Test; Presentation modality; Pictorial superiority effect; Development; Sex.

Working memory is believed to be an elementary cognitive function, that is positively related to other cognitive domains, such as language, reasoning, and attention, and academic success in terms of, for example, reading and mathematics (Alloway, Gathercole, Kirkwood, & Elliot,

No potential conflict of interest was reported by the authors.

Address correspondence to Celeste Meijs, c/o Welten Institute, Open University, P.O. Box 2960, 6401 DL Heerlen, The Netherlands. E-mail: celeste.meijs@ou.nl

Present address of Celeste Meijs is Welten Institute, Research Centre for Learning, Teaching and Technology, Open University, The Netherlands.

2009; Constantinidou, Danos, Nelson, & Baker, 2011; Hitch, Halliday, Schaafstal, & Heffernan, 1991). In terms of the tailoring of instructions and the presentation of information that needs to be learned in, for instance, schools and other educational contexts, it is essential to understand the development of working memory (Cherry et al., 2008; Tam, Jarrold, Baddeley, & Sabatos-DeVito, 2010).

Baddeley (2003, 2012) has defined working memory as a relatively loose theoretical framework which consists of four components: a central executive, the phonological loop, the visuospatial sketchpad, and the episodic buffer (Baddeley, 2012). The central executive is responsible for the attentional control of working memory, coordinating (a) the retrieval of information from long-term memory and (b) the phonological loop and the visuospatial sketchpad (Nevo & Breznitz, 2013). Hence, the phonological loop and the visuospatial sketchpad are called the subsidiary slave systems and are modality specific. The visuospatial sketchpad integrates visual and spatial information into a unified ("integrated") representation. This information can temporarily be stored and manipulated (Baddeley, 2003; Nevo & Breznitz, 2013). The phonological loop is a modular system that contains two subcomponents: a store to briefly hold phonological (verbal) information and a (sub)vocal rehearsal system that (a) is used to maintain information and (b) registers or translates visual information that can be "named" into a phonological code for storage in the phonological store (Baddeley, 2003, 2012; Palmer, 2000). The (sub)vocal rehearsal system is involved in the "pictorial superiority effect", namely that, in adults, nameable items are remembered better if presented in the form of a picture (visual, e.g., a picture of a pineapple) as compared to auditory presentation (the spoken word "pineapple"). On the one hand, auditory information is processed automatically, perhaps even obligatory, in a phonological code (Baddeley, 2003, 2012). Thus, auditory information processing is unimodal. On the other hand, nameable visual (or pictorial) information is in general processed via multimodalities. Pictorial information is recoded to a phonological code and hence processed in the visuospatial sketchpad (visual information of the picture) and again via the phonological loop (phonological information of the word belonging to the picture). This dual-loop activation increases the chance of remembering (and learning) the information, compared to the unimodal activation of auditory presentation (spoken words).

For textual information (e.g., typed words), sequences of typed letters or digits are recalled based on the stored acoustic or phonological stored information in experienced readers (Baddeley, 2003; Halliday, Hitch, Lennon, & Pettipher, 1990). The visual information of the printed letters or digits will not contribute to a better recall because the visual information of the printed letters or digits will not be processed in depth in the visuospatial sketchpad. Thus, typed words are, similar to the auditory presentation of words, subject to unimodal processing in experienced readers. Finally, the episodic buffer combines visual and verbal (or phonological) codes and links these codes to multidimensional presentations in long-term memory and is concerned with the storage of information (Baddeley, 2003).

As mentioned above, even though adults process nameable pictures in a phonological code, visual coding still takes place (Palmer, 2000). Visual codes are complemented by phonological codes rather than replaced by them (Kemps, De Rammelaere, & Desmet, 2000). Thus, both visual and phonological memory codes are present in the processing of nameable visual information, but the phonological coding masks the visual coding (Hitch, Woodin, & Baker, 1989) and overshadows what is referred to as imaginal encoding (which is visual in nature; Halliday et al., 1990). This pattern of processing suggests a dual coding of phonological and nameable visual information, which is also the basic assumption of another theory that is often used to explain the pictorial superiority effect—dual coding theory (DCT; Paivio, 1969; Thompson & Paivio, 1994). According to DCT, the number of representations in the brain that are evoked by the presented material determine how well the information will be remembered. The two codes that are evoked by the pictorial presentation of nameable visual items increase the chance of recall, as compared to the one code evoked by auditory presentation (Paivio, 1969). The difference between both theories—Baddeley's (2003) model and the DCT model—is that, in the working memory model, the emphasis is on the process of recoding nameable visual information into a phonological code via the (sub)vocal rehearsal system (which leads to dual processing), whereas in DCT, the emphasis is on dual processing, which is linked to inner speech and speech rate.

Thus, in adults, the optimal presentation modality for learning verbal/phonological information (such as words) would therefore be visual, for instance, in the form of a picture. However, for children, the optimal presentation modality is less clear because of inconsistencies in the literature on the development of working memory. On the one hand, studies indicate that working memory is functional at an adult level as early as 6 years of age (Alloway, Gathercole, Willis, & Adams, 2004), whereas, on the other hand, studies indicate that working memory components are still improving between the ages of 7 and 11 years (Halliday et al., 1990; Magimairaj & Montgomery, 2012). Nonetheless, the majority of studies claim that the (sub)vocal rehearsal process in the phonological loop becomes active around the age of 7 years (Baddeley & Hitch, 2000; Cowan, 1998; Kemps et al., 2000; Tam et al., 2010). Before this, it is assumed that the processing of visual information (pictures) takes place solely in a visuospatial code (Halliday et al., 1990; Henry, Turner, Smith, & Leather, 2000) because the system needed for recoding (or "labeling") the nameable visual information into words (phonological) is not functional yet. Hitch et al. (1991) link spontaneous labeling to the activation of the articulatory loop (which Baddeley refers to as the phonological loop) and to "inner speech", which is, according to Hitch et al., not fully developed in 5-year-old children. However, Palmer (2000) indicates that the developmental pattern of the (sub)vocal rehearsal system is more nuanced. In her study, she found that the majority of the children (82%) showed a shift from the use of no coding strategies of pictures during recall to the use of visual coding strategies (i.e., visual processing of the pictures), followed by the use of both visual and verbal coding strategies (i.e., labeling of the pictures) and later to the use of predominantly adult-like verbal coding strategies during recall. This developmental pattern has no onset at a particular age and varies from child to child. This is in line with the findings of Kemps et al. (2000), who suggest that some 5-year-olds recode nameable visual information in words whereas other 5-year-olds do not, and that (almost) all children spontaneously and consistently recode visual information around the age of 8 to 9 years. Taken together, the process of recoding of nameable visual information (e.g., pictorial) into a phonological code by the (sub)vocal rehearsal system might become active at some point between the ages of 5 and 8 and, therefore, the pictorial superiority effect would manifest between those ages. According to Whitehouse, Maybery, and Durkin (2006), an increase in performance differences with age between information processed in two codes (nameable visual information) and one code (in their study a textual presentation) is caused by an increase in the functionality of inner speech when children become older. This suggests an increase in the pictorial superiority effect with age.

Most studies into the optimal presentation modality make use of tests based on a working memory paradigm. Commonly used tests of verbal working memory are those which measure digit span, word span, or letter span, while popular visual or spatial array tests include the Corsi block test and the tap-test. These working memory tests have several features. First, information is presented either verbally (i.e., digits, words, or numbers) or visuospatially (i.e., patterns of positions on an array, assumed to be not nameable). Second, the information is presented in a series which increases in length (and hence in level of difficulty) that contains differential information for each series (i.e., different sequences of digits). Third, recall is immediate, in the same (fixed) order (i.e., short-term memory or simple span) or after manipulation of the information (e.g., recall in reversed order or while simultaneously processing related information such as the meaning of sentences. i.e., complex memory span; for more information on working memory tests, see Alloway et al., 2004; Baddeley, 2003; Kemps et al., 2000; Magimairaj & Montgomery, 2012). Unfortunately, even though these paradigms are excellent tests for the assessment of working memory, the paradigms differ from how information is presented and learned in educational contexts such as schools. Features of learning in the classroom include repeated presentation of the same information, test of recall after a prolonged period of time, and no fixed order of recall. Thus, the ecological validity of working memory paradigms is assumed to be low, if studying learning in educational contexts is the aim.

Therefore, in the current study, we chose to use another paradigm that more closely resembles learning in educational context to study developmental differences in the presentation modality of information that is to be learned: the Verbal Learning Test (VLT). The VLT features five repeated presentations of a list of words (trials), with an immediate and delayed recall and recognition condition, sometimes supplemented with the presentation of a list with distractor words following the fifth presentation (see the Methods section below). The advantage of using the VLT paradigm is that it is possible to study several aspects of verbal learning that also occur during learning in an educational context. First, it gives an indication of the verbal learning capacity, in which the progress of learning with repeated presentations can be monitored and overall verbal learning can be calculated by summing the performances on each presentation. This measure reflects acquisition and learning (Vakil, Greenstein, & Blachstein, 2010). Unsworth and Engle (2006) describe that the first presentations of a VLT resemble the procedure of a supra-span measure (i.e., the presentation of a list of items above the memory span). Second, delayed verbal learning capacity and verbal learning recognition capacity can be studied when the information is recalled after a period of approximately 15 to 20 minutes. Delayed verbal learning capacity reflects retention and forgetting. Delayed verbal recognition gives insights into whether a low delayed verbal capacity reflects a problem in retrieval or whether problems occur during encoding and storage. Recognition is a valuable measure in younger children because it gives a more precise estimate of the memory functions in this age group (Vakil et al., 2010). Encoding and retrieval are two processes that are used to retrieve information from what Unsworth, Fukuda, Awh, and Vogel (2014) refer to as “secondary memory”, which can be perceived as long-term memory. These features of verbal learning in a VLT are more commonly seen in educational contexts such as schools as compared to the working memory paradigms and, therefore, verbal learning in a VLT will be the focus of the current study.

Although, numerous studies have made use of a VLT, most developmental studies have had a narrow focus. For instance, most developmental studies that include a VLT only use auditory presentation (Bjorklund & Douglas, 1997; Lezak, Howieson, & Loring, 2004; Schneider & Pressley, 1997; Van Den Burg & Kingma, 1999). Hence, much less is known about the effects of other presentation modalities on verbal learning performance, such as a pictorial or textual presentation. Halliday et al. (1990, p.36) state: "In general, it seems that mode of presentation has been neglected in developmental studies as a factor of both methodological and theoretical importance". Even though this statement was made in 1990, few studies have been conducted into the developmental differences and the relation with presentation modality in a VLT since, and have been inconclusive or have had a narrow focus. This gap in the knowledge has been referred to as surprising (Whitehouse et al., 2006), especially since reported studies have shown that a pictorial presentation (as compared to an auditory presentation) often leads to better performance in a VLT in various populations, ranging from college students to the elderly (Cherry et al., 2008; Constantinidou & Baker, 2002; Lezak et al., 2004; Toglia, Hinman, Dayton, & Catalano, 1997). Whitehouse et al. (2006) performed one of the few VLT studies in which a presentation with typed words was compared to a pictorial presentation, and they reported an increasing pictorial superiority effect with age. Unfortunately, this study was performed exclusively with boys of 7 to 15 years of age. A study that compared (a) a pictorial presentation with (b) an auditory presentation and (c) a pictorial-auditory combined presentation reported better performance in the pictorial and pictorial-auditory presentations as compared to the auditory presentation (Constantinidou et al., 2011). However, this study only included children of 7 to 13 years of age. We hypothesize that, regarding age, the most interesting findings will be found below the age of 7 years. Additionally, these studies did not investigate interactions between the presentation modality and sex.

Another factor that needs more attention is the potential difference of the influence of presentation modality related to sex, which is, to our knowledge, yet to be studied. If sex differences were studied, these focused on VLT performance independent of presentation modality. Studies that report sex differences indicate that girls outperform boys in overall VLT performance, which has been ascribed to better memory strategy use by girls (Cox & Waters, 1986). However, some studies reported only marginally better verbal learning performance by girls (Van Den Burg & Kingma, 1999) or reported that sex differences (in adults) were too subtle to notice (Kramer, Yaffe, Lengenfelder, & Delis, 2003), while other studies reported no sex differences (Alloway, Gathercole, & Pickering, 2006; Gathercole, Pickering, Ambridge, & Wearing, 2004). However, even though the findings in literature are not consistent regarding sex differences, the majority of the findings seem to indicate that girls perform (marginally) better in VLTs than boys. Therefore, we hypothesize that girls will outperform boys for all measures of the VLT. However, information on the interaction between presentation modality and sex is lacking, therefore we explored this in the current study. We hypothesized that there would be no interaction between presentation modality and sex.

Taken together, there are no published studies that, simultaneously, focus on: (a) the comparison of three presentation modalities: pictorial, auditory, and textual, (b) children aged 5 to 16 years, and (c) both boys and girls in a VLT paradigm. Our study is the first to fill this gap.

The first aim was to specifically study the development of the pictorial superiority effect through the inclusion of children in which the articulatory control process is assumed to be not yet active (e.g., children below the age of 7). Therefore, in the first wave of the study, children aged 5 to 15 (ranging from grade 2, primary school [P2] to grade 2, secondary school [S2]) performed a pictorial VLT (PVLT) or an auditory VLT (AVLT). The first hypothesis was that children below the age of 7 would show no differences in performance in the PVLT and the AVLT, whereas children above the age of 7 would show a better performance in the PVLT than the AVLT. The second hypothesis was that girls would outperform boys in both the AVLT and the PVLT because of the assumed better verbal learning capacity of girls. Thus no interaction effects between presentation modality and sex were expected.

The second aim of the study was to specifically study the effects of three presentation modalities (PVLT, AVLT, and a textual VLT [TVLT]) in children aged 9 to 16 (ranging from grade 5, primary school [P5], to grade 3, secondary school [S3]). For this part of the study, we selected children in grades 5 and higher because reading should be mastered at a sufficient level to be able to perform the TVLT. This was studied during the second wave of the study, one year later, in which the same children completed a PVLT, AVLT, or TVLT. We hypothesized that performance in the PVLT would be better than the performance in the AVLT and the TVLT in all grades, for the reasons mentioned above.

METHOD

Participants and Procedure

The current study was part of a large longitudinal study into cognitive development. In total, the longitudinal study consisted of three waves, of which we used data from the first and second wave. In the first wave, participants were recruited from 29 regular primary and secondary schools in the city of Maastricht, The Netherlands and its surroundings. The parents/caregivers (referred to as *caregivers* hereafter) of the children enrolled in primary school grades 2, 4, 6, and 8 (referred to as P2, P4, P6, and P8 respectively) and those enrolled in secondary school grades 1 and 2 (ranging from pre-vocational secondary education to pre-university education, referred to as S1 and S2 respectively) received an information package via their schools describing the purpose of the study, a request to participate, a questionnaire, a form to give (informed) consent for the child to participate, and a stamped addressed envelope.

Of the 1086 caregivers who replied, 892 (82%) gave consent for their child to participate. Children who met the following criteria were eligible for participation: the child (1) had not repeated and/or skipped a grade, (2) was of Dutch nationality (in view of assumed fluency in Dutch), and (3) did not take medication that might influence cognitive performance. The children were randomly selected by school grade level and sex per grade. In total, 431 children were included in this first wave. However, not all children had reliable scores (see the subsection on missing data and extreme values for details) for the test of interest (VLT) and this resulted in the inclusion of 408 children in wave 1 of the study (see Table 1 for the demographics). In the second wave, one year later, the caregivers of the previously tested children were contacted and asked to participate again. Of these, 336 caregivers responded, and 333 gave consent for their child to participate. After the selection of the children who had not repeated or skipped a grade between the first and second wave (n = 8), 325 children were scheduled for testing. Only 313 children were tested, because (a) some children refused to participate (primarily children in secondary school) or (b) it was not possible to test them within the scheduled time frame (e.g., because of illness or interfering school activities). In the data used from the second wave, only children from grades 5 and higher were included (n = 237, see Table 1 for the demographics) because of the mastery of reading required to perform the test of interest (i.e., the TVLT).

Table 1 Demographics of the Included Children.

			Sex		Presentation Modality		
Grade	*n*	Age(*SD*)	Boys	Girls	PVLT	AVLT	WVLT
Wave 1							
P2	69	6.31(0.33)	34	35	37	32	-
P4	67	8.37(0.44)	36	31	34	33	-
P6	78	10.35(0.36)	33	45	37	41	-
P8	66	12.30(0.41)	31	35	35	31	-
S1	70	13.37(0.38)	39	31	34	36	-
S2	58	14.44(0.33)	31	27	24	34	-
Total	408		204	204	201	207	-
Wave 2							
P5	57	9.38(0.41)	29	28	23	18	16
P7	56	11.43(0.35)	20	36	18	18	20
S1	43	13.61(0.87)	20	23	15	14	14
S2	49	14.37(0.33)	29	20	15	14	20
S3	32	15.70(1.01)	20	12	11	14	7
Total	237		118	119	82	78	77

In both waves, caregivers filled in a questionnaire giving background information on their child, and all children completed a battery of several neuropsychological tests, including the test of interest: a VLT. Well-trained psychological assistants (i.e., undergraduate psychology students) tested the children in separate rooms at their schools. Testing of the complete battery of tests took approximately 1.5 hours, administration of the VLT lasted approximately 20 minutes (retention time not included). All tests were administered in the same order for each child. The Ethics Committee of the Faculty of Psychology & Neuroscience of Maastricht University approved the study protocol.

Instruments

A VLT is often used in clinical settings and memory research, and is one of the most sensitive verbal memory tests for the assessment of verbal learning and memory abilities (Constantinidou & Baker, 2002). Its test-retest reliability is reported to be high (Lezak, 1995; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2005). Three parallel versions of the VLT were constructed for the purpose of the current study: a word list with a pictorial, verbal, or textual presentation modality. All words included in these different VLT versions were equivalent in terms of their frequency of use in the Dutch language (Linschoten, 1963) and their imageability (Van Loon-Vervoorn, 1989). The selected words were also ones which are acquired before the age of 6 years (Van der Elst et al., 2005). Special attention was paid to the dispersion of words that might evoke automatic clustering within the three lists (e.g., avoidance of clusters of words which start with the same letter). These three (statistically proved) parallel versions were designed to prevent learning effects because of the longitudinal set-up of the study. The wordlists were composed of semantically unrelated words, which would limit clustering of words based on meaning.

The test included an immediate recall, in which 15 words were presented five times (in the same order). These presentations were called “trials”. After each presentation, the children had to recall (verbally) as many words as possible, in the order they preferred. They were given no specific instructions on how to remember the words because we were interested in what the children would spontaneously do to remember as many words as possible. After 15 to 20 minutes (the retention time), in which no memory interfering test was administered, there was a delayed recall condition in which the children had to recall as many words as possible without a presentation. Then, a recognition condition followed, in which 30 words, including the 15 words of the presented list, were presented to the children. They had to reply with "Yes" or "No" to indicate whether or not the word belonged to the previously presented wordlist. This procedure is identical to standard VLT assessment.

The procedure of the administration of the test was the same for all modalities. All words were presented via a computer at a rate of one item every two seconds and the presentation time of the pictures and the typed words equaled the mean presentation duration of the auditory presentation. For the auditory presentation, the words were recorded and pronounced by a male voice.

In the first wave, half of the children performed the test in which the words were presented as concrete line-drawings (PVLT) and the other half performed the test in which the words were listed via an auditory presentation (AVLT). They were assigned to the presentation modalities via stratified sampling; within the strata "grade and sex", they were randomly assigned to the two conditions. In the second wave, from grade P5 onwards, one third of the children performed the PVLT, one third performed the AVLT and one third performed the TVLT. One third of the children that performed the PVLT in wave 1 and one third of the children that performed the AVLT in wave 1 performed the TVLT in wave 2. The children who did not perform the TVLT performed the VLT in the same modality as they did in wave 1.

Statistics

Missing Data and Extreme Values. Before data analysis, extreme values and missing data were assessed. First, the verbal learning measures were checked for extreme values, defined as values minimally three times the interquartile distance above the 75th percentile or below the 25th percentile (Huizingh, 2002), except for verbal learning recognition memory because this measure was not normally distributed. In wave 1, no extreme values were detected. In wave 2, a total of 3 children had extreme scores on trial 1 of the verbal learning capacity measure and 1 child had an extreme score on trial 4 of the verbal learning capacity measure. These scores were recoded as "missing data" for that particular measure only. However, as a consequence, the overall verbal capacity measure (i.e., a composition score including trial 1 and trial 4) was not calculated for these 4 children.

Second, we checked for missing data. A measure was rated as "missing" if the test was not administered (for instance because of time constraints), if the score was an extreme value (as defined above), or if the administration was unreliable (indicated by the psychologist administering the tests). Unreliability could be due to several factors, including fatigue of the child, failure of equipment, or interfering environmental factors such as suddenly excessive background noise. In wave 1, the data of 408 children was reliable for the verbal learning capacity measures. A total of 6 children had unreliable scores for the delayed verbal learning capacity and 14 children had unreliable scores for the verbal learning recognition capacity. However, the data of these children were included in the analyses of the verbal learning capacity measures (i.e., trials 1–5) and the overall verbal learning capacity measure (i.e., the sum of trials 1–5)—outcome measures that did not include the delayed recall or delayed recognition scores.

In wave 2, the data of 237 children was included in the immediate recall. No missing values were found, other than the ones defined as missing because of extreme values (as detailed above). Missing data were not replaced because less than 5% of the data were missing per measure (Croy & Novins, 2005).

Variables and Analyses. The first aim was to study the development of the pictorial superiority effect and whether this differs for boys and girls. To study this in detail, we analyzed the performances for the verbal learning measures—verbal learning capacity (trials 1–5), overall verbal learning capacity (sum of trials 1–5), delayed verbal learning capacity, verbal learning recognition capacity (sum of correctly recognized and correctly rejected words)—per grade separately, because we expected that the differences would be most pronounced in the youngest grades. Testing over all grades in one analysis would hence not be sufficiently sensitive to disclose the specific differences between grade P2 and the older grades as a group. Testing per grade was achieved by dividing the dataset by means of a split-file per grade and analyzing the differences in performance for the verbal learning measures between the PVLT and the AVLT. Sex was included as additional independent variable to study sex effects on verbal learning and whether presentation modality differentially affected boys and girls.

General linear model (GLM) univariate analyses were performed with presentation modality (PVLT, AVLT) and sex (boys, girls) as independent variables and the measures of the VLT as dependent variables. The α-level was set to .01 to correct for type-I errors due to multiple testing. We preferred GLM analyses over regression analyses because we were specifically interested in differences per grade and we had not aimed to predict performance. Since performance scores for the verbal learning recognition capacity were not normally distributed (due to a ceiling effect), non-parametric tests were performed—two independent samples tests (Mann Whitney U; Chi-square analyses) with presentation modality (PVLT, AVLT) and sex (boy, girl) as independent variables and verbal learning recognition capacity as a dependent variable.

The second aim was to study the effect of a textual presentation modality, additional to the pictorial and auditory presentation of the words on the measures of a VLT, and the possible differences with increasing age. To study this in detail, we analyzed the performance for the verbal learning measures (verbal learning capacity, overall verbal learning capacity, delayed verbal learning capacity, and verbal learning recognition capacity) per grade separately by means of a split-file for grade. Due to lower cell counts, non-parametric tests were performed—two independent samples tests (Mann Whitney U tests; Chi-square analyses) with presentation modality (PVLT and AVLT, PVLT and TVLT, and AVLT and TVLT) as (paired) grouping variables. Two independent samples tests (instead of several independent samples tests) were used because this enabled us to study the differences on the level of one presentation modality against another. If several independent samples tests had been used (comparing all three presentation modalities at once), then it would not have been possible to distinguish between the effects of the presentation modalities per presentation modality pair.

RESULTS

Aim 1

Means and standard deviations (*SD*s) of the measures of the VLT per modality and per grade for wave 1 are summarized in Table 2. The results of the effects of presentation modality on the GLM univariate analyses can be seen in Table 3. Significant effects of presentation modality were found in grade P2 for recognition (Mann Whitney U = 365.00, $Z = -2.77$, $p = .006$; not in Table 3). In grade P4 and higher, overall, significant results were found for trial 1, total recall, and delayed recall, and some for trial 2 in the highest grades. In all cases, the number of words recalled and/or recognized in the PVLT was higher than the number of words recalled and/or recognized in the AVLT. Figure 1 displays these results for trial 1, total recall and delayed recall.

No interaction effects of presentation modality and sex were found. However, a main effect of sex was found for delayed recall for children in grade P2 ($F(1, 63) = 7.38$; $p = .009$), which indicated that the girls outperformed the boys.

Aim 2

Means and *SD*s of the measures of the VLT, per modality and per grade for wave 2, are summarized in Table 4. The results of the independent samples analyses can be seen in Table 5. Significant effects of presentation modality were found for grade P5 on

Table 2 Means and *SD*s of the Measures of the VLT, Per Modality, Per Grade for Wave 1.

	P2	P4	P6	P8	S1	S2
Trial 1						
PVLT	4.84(1.62)	6.62(2.00)	7.62(1.57)	8.31(1.71)	8.59(1.72)	8.29(1.40)
AVLT	4.47(1.41)	4.94(1.25)	5.83(1.63)	6.94(1.91)	6.86(1.51)	7.15(1.52)
Trial 2						
PVLT	6.11(1.93)	8.00(2.66)	9.73(2.49)	10.29(1.93)	10.91(2.35)	11.17(1.43)
AVLT	5.88(1.64)	6.64(2.09)	8.46(2.00)	9.06(2.25)	9.53(1.92)	9.32(1.70)
Trial 3						
PVLT	6.62(2.30)	9.03(2.59)	10.97(1.88)	11.09(2.15)	11.88(2.09)	11.25(1.54)
AVLT	5.53(2.45)	8.06(2.47)	9.85(1.98)	10.52(2.05)	10.72(1.97)	11.21(1.90)
Trial 4						
PVLT	7.30(2.59)	9.76(2.63)	11.68(2.08)	11.63(1.66)	12.00(2.03)	11.83(1.34)
AVLT	6.97(2.89)	8.52(3.18)	10.37(1.67)	10.65(1.72)	11.50(1.89)	11.09(2.05)
Trial 5						
PVLT	7.19(3.12)	10.03(2.08)	11.84(1.48)	12.14(1.75)	12.56(1.58)	12.75(1.67)
AVLT	7.56(2.54)	9.42(2.46)	10.34(2.37)	11.42(2.00)	11.94(1.90)	11.56(1.86)
Total recall						
PVLT	32.05(9.064)	43.44(9.535)	51.84(6.384)	53.46(6.621)	55.94(8.004)	55.29(5.154)
AVLT	30.41(8.691)	37.58(9.404)	44.85(7.475)	48.58(7.150)	50.56(7.666)	50.32(6.709)
Delayed recall						
PVLT	7.86(3.09)	10.41(1.88)	11.47(1.81)	11.85(1.89)	11.68(2.42)	12.04(1.71)
AVLT	6.87(2.28)	8.52(2.39)	9.83(2.46)	10.47(2.21)	10.77(2.65)	10.44(2.39)
Recognition						
PVLT	29.77(0.55)	29.76(0.50)	29.74(0.89)	29.91(0.29)	29.85(0.56)	30(0.00)
AVLT	29.06(1.48)	29.31(1.35)	29.66(0.67)	29.73(0.58)	29.77(0.59)	29.65(0.77)

Table 3 Results of the GLM Univariate Analyses with Modality as the Dependent Variable (PVLT vs AVLT) and the Measures of the VLT as Dependent Variables, with a Split File for Grade.

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Total recall	Delayed recall
K							
F(1,65)	1.04	0.18	3.39	0.19	0.44	0.47	1.71
p	.314	.673	.070	.668	.523	.469	.195
PES	.016	.003	.050	.003	.006	.007	.026
P4							
*F(*1,63)	16.58	4.94	2.27	3.15	1.06	6.13	12.40
p	< .001*	.030	.135	.081	.306	.016	.001*
PES	.208	.073	.035	.048	.017	.089	.164
P6							
F(1,74)	22.87	6.48	6.41	9.85	11.11	19.86	12.15
p	< .001*	.013	.013	.002*	.001*	< .001*	.001*
PES	.236	.081	.080	.117	.131	.212	.143
P8							
F(1,62)	9.25	5.23	1.12	5.48	2.14	7.85	7.03
p	.003*	.026	.294	.022	.148	.007*	.010*
PES	.130	.078	.018	.081	.033	.112	.105
S1							
F(1,66)	20.84	8.77	5.61	1.17	2.26	8.79	2.13
p	< .001*	.004*	.021	.283	.138	.004*	.149
PES	.240	.117	.078	.017	.033	.118	.032
S2							
F(1,54)	8.62	18.43	0.00	2.16	5.76	8.90	7.43
p	.005*	< .001*	.945	.148	.020	.004*	.009*
PES	.018	.254	.000	.038	.096	.042	.121

Note. * = significant. For all significant findings, PVLT > AVLT.

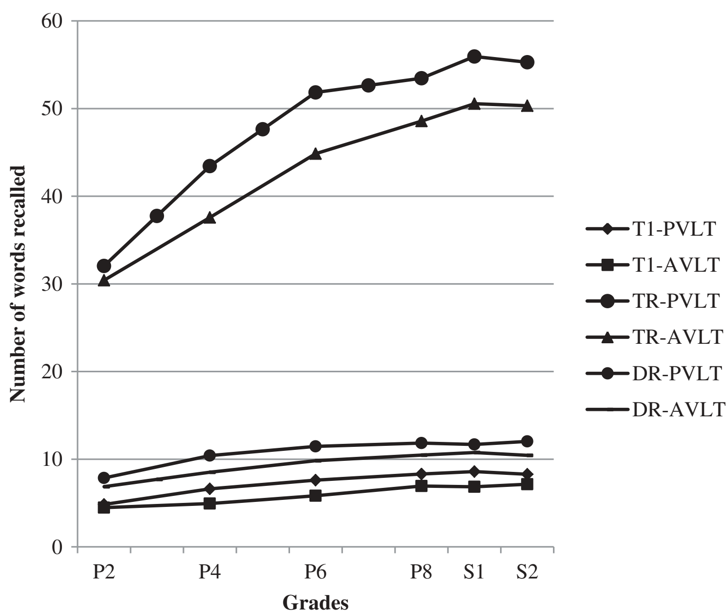

Figure 1 Number of words recalled, per grade, per modality for performance in trial 1, total recall and delayed recall, in wave 1.

Table 4 Means and *SD*s of the Measures of the VLT, Per Modality, Per Grade for Wave 2.

	P5	P7	S1	S2	S3
Trial 1					
PVLT	7.35(1.94)	9.22(1.52)	8.67(1.34)	9.60(1.80)	8.64(1.69)
AVLT	5.94(1.30)	6.22(1.59)	7.50(0.94)	6.69(1.32)	7.50(1.45)
WVLT	6.50(1.63)	6.95(1.32)	7.71(1.68)	7.75(2.00)	8.71(2.29)
Trial 2					
PVLT	9.65(2.08)	11.72(1.56)	11.60(1.45)	11.73(1.58)	11.00(2.28)
AVLT	8.39(1.19)	9.06(1.39)	10.64(2.20)	10.29(1.59)	11.07(1.64)
WVLT	8.81(2.04)	9.15(2.11)	10.29(1.64)	10.55(2.01)	11.29(2.69)
Trial 3					
PVLT	10.43(1.78)	11.72(1.53)	12.87(1.06)	12.33(1.59)	12.00(2.24)
AVLT	9.50(1.04)	10.67(1.94)	12.14(1.96)	12.00(1.84)	11.93(1.94)
WVLT	9.75(1.73)	11.30(1.720)	11.57(2.06)	11.55(1.79)	11.43(0.98)
Trial 4					
PVLT	11.35(1.58)	11.72(2.08)	12.27(1.87)	12.60(1.35)	12.36(1.91)
AVLT	10.72(1.49)	11.11(1.57)	12.21(2.15)	12.43(1.60)	12.86(1.75)
WVLT	11.40(2.063)	11.10(1.59)	12.00(1.36)	12.70(1.69)	12.00(1.63)
Trial 5					
PVLT	11.52(1.62)	12.50(1.92)	13.47(0.83)	12.67(1.50)	12.64(1.21)
AVLT	11.00(1.68)	11.22(2.39)	12.93(1.82)	12.64(1.55)	13.07(1.38)
WVLT	10.94(1.69)	11.55(1.19)	12.36(1.34)	12.75(2.02)	12.57(1.40)
Total words					
PVLT	50.17(7.011)	56.89(5.519)	58.87(4.75)	58.93(5.98)	56.64(8.19)
AVLT	45.78(3.949)	48.28(7.521)	55.43(7.58)	53.54(5.97)	55.25(5.86)
WVLT	47.8(7.292)	50.00(6.130)	53.93(5.55)	55.30(7.78)	56.00(7.75)
Delayed recall					
PVLT	10.48(2.21)	12.28(1.708)	12.13(1.77)	12.33(1.80)	11.82(2.04)
AVLT	9.17(0.79)	10.44(2.332)	10.79(2.69)	11.36(1.82)	11.57(1.78)
WVLT	9.5(1.932)	9.60(2.234)	11.29(1.73)	11.45(2.06)	11.43(1.81)
Recognition					
PVLT	29.96(0.209)	29.94(0.236)	29.93(0.26)	29.93(0.258)	30.00(0.000)
AVLT	29.22(1.11)	28.94(1.697)	29.64(0.63)	29.79(0.426)	30.00(0.000)
WVLT	29.87(0.34)	29.85(0.366)	29.93(0.27)	29.85(0.366)	29.43(1.512)

recognition and for some measures in grades P7 and S2. Overall, if differences were found, performance in the PVLT was better than that of the AVLT and/or TVLT.

DISCUSSION

The findings of the current study show that there is no pictorial superiority effect in children below the age of 7. This conclusion was based on the finding that no differences between PVLT and AVLT performance were found for almost all outcome measures in children younger than 7 years of age (i.e., grade P2). However, one difference was found for verbal learning recognition capacity, in which pictures were better recognized than the auditory presentation of words. The difference between recognition capacity and the other measures of the VLT is that during recognition there is no active recall or active search in memory for the words. Therefore, this may indicate that this measure reflects storage and encoding, which is a purer measure of storage than other measures that also include retrieval (Vakil et al., 2010). However, because recognition is less complex, it is also less demanding—and this leads to ceiling effects if (overall) verbal learning capacity increases (as in older children). Still, the significant difference found for recognition indicates that, even in grade P2, pictures are stored better than the auditory presentation of words. It is possible that the active retrieval and memory search for the words hampers pictorial superiority, at least in children under 7 years of age. The working memory model is developed by means of tests that use measures which include retrieval processes. Therefore, the fact that we found no differences in performance between the PVLT and the AVLT for the measures of the VLT that include retrieval processes is in line with the working memory model (Baddeley, 2012; Baddeley & Hitch, 2000; Cowan, 1998; Kemps et al., 2000; Palmer, 2000), and more specifically with the development of the (sub)vocal rehearsal system in the phonological loop. The loop is used for the recoding (or "labeling") of the pictorial information into a phonological code for further processing and is assumed to become functional at the age of approximately 7 years. Not only are these findings in line with the working memory model, but the DCT explains them in terms of the development of inner speech and the corresponding speech rate (Magimairaj & Montgomery, 2012; Tam et al., 2010; Whitehouse et al., 2006), which are assumed to reflect the same underlying mechanism as the working memory model ascribes to the (sub)vocal rehearsal system. However, actively retrieving the words from memory might be such an effortful process that it overshadows the actual storage of the words, leaving the subtle differences between the presentation modalities undetected. Thus, the assumption that the (sub)vocal rehearsal process is already in use by some children (Kemps et al., 2000) and the statement that there is no concrete onset age at which the (sub)vocal rehearsal process becomes active and that this might be before the age of 7 years for some children (Palmer, 2000) is confirmed by the current findings.

Additionally, the findings of the current study show that there is a pictorial superiority effect in children above the age of 7, but not for all measures of the VLT. This conclusion is based on the finding that in children above the age of 7 (P4 and higher), overall, PVLT performance was better than AVLT performance for those measures that are believed to be most strongly dependent on the recoding of the pictures—namely the first (and, less pronounced, the second) presentation (trials 1 and 2)—and the measures that are related to the first presentation(s)—namely overall verbal learning capacity (sum of trials 1–5) and delayed verbal learning capacity. Thus, the recoding of the pictures evokes better recall of the pictures as compared to the words listed via auditory presentation. These findings also suggest the involvement of the (sub)vocal rehearsal system in children above the age of 7, but also that this involvement is largest in the first trials of the VLT. The finding that the pictorial superiority disappears on the third, fourth and fifth presentations (except for grade P6, for more information see below) suggests that after the first presentation, children of ages 7 and older use other memory processes or context information to remember and recall the information. Hence, the current study confirms that the VLT taps into several types of learning mechanism. Baddeley (2003) describes that a VLT which has repeated presentations might invoke processes that are involved in long-term memory. The central executive and the episodic buffer might be involved when links are made between the words and the semantic meaning of those words in long-term memory. This process is assumed to emerge when the items have been recoded or labeled, which is after the first and second presentations (i.e., trials 3–5). The results of the current study suggest that presentation modality does not influence these processes. This is in line with the working memory model which states that the central executive and episodic buffer are not modality specific. In the highest grades (S1 and S2), PVLT performance in trial 2 was better than AVLT performance, whereas no effects were found in trial 2 for the younger children. This finding is in line with the findings of Whitehouse et al. (2006) stating that the pictorial superiority effect increases with age.

Remarkably, in grade P6, PVLT performance was better for trials 4 and 5. This finding might be related to that specific group of children (since the non-significant findings for trials 2 and 3 are nearing significance) and we will explore this finding in more detail by comparing these findings with the findings of this age group in wave 2, when these children attended grade P7.

Also, the findings in wave 2 indicate that there is no difference in performance if words are presented in pictorial, auditory or textual form in children P5 and older. This is not in line with the DCT (Paivio, 1969) and the study conducted by Whitehouse et al. (2006) that reported better performance in the PVLT as compared to the AVLT and the TVLT for all children. Additionally, it contradicts our findings in wave 1. However, we acknowledge that our groups might be too small to detect differences and this seems to be the case if we inspect the results of the analyses (see Table 5). In trial 1, differences between PVLT and AVLT performances might have been significant if the groups were larger, as was the case in wave 1. However, other non-significant performance (in the other trials and for the auditory and textual presentation condition) do not meet these findings. Nevertheless, the findings of the current study suggest that there is no difference in the processing of words presented in auditory or textual form. This may be explained by means of both the working memory model—the typed words do not need to be labeled or recoded because they have a phonological code—and the DCT—both auditory presented words and typed words are processed in one modality only, namely auditory.

Some significant differences due to presentation modality were found. Most of these differences were found for children in P7. As described above, in wave 1, these children, who attended P6 at that time, showed a different pattern of influences of presentation modality in that wave also. Thus, we assume that there might be other factors that influence the performances on the VLT of these children. We inspected the *SD*s of the performances per grade, per trial and per modality, but no deviancies were found (see Tables 2 and 4). Thus, no exceptional higher or lower performing children were part of respectively the PVLT or the AVLT group (i.e., no extreme values). If we compare the learning curve, by inspecting the gradual increase in performance over the trials per modality for grade P6 (wave 1), we see a gradual increase for both modalities with a steeper increase in performance in the first trials and a leveling off in the later trials (see Table 2). There were no differential learning curves for the PVLT and the AVLT. Compared to other grades, again, there were no deviancies. However, if we inspect the plots of the learning curves in wave 2 (see Table 4), we see that the children in P7 who performed the PVLT had higher scores than expected, namely they performed better than the children in S1. Additionally, boys and girls were distributed evenly. Thus, it seems likely that the children who performed the PVLT in P6 in wave 1 and P7 in wave 2 were, regardless of presentation modality and sex, as a group, better-performing children.

The children in S2 performed better in the PVLT than in the AVLT and the TVLT in trial 1, and performed better in the PVLT than in the AVLT on total recall. In wave 1, in which these children attended S1, we saw a similar pattern, namely better PVLT performance than AVLT performance in trial 2. We concluded that the pictorial superiority effect increases with age, which these findings underscore. However, the results of the children

Table 5 Results of the Two Independent Samples Tests with PVLT vs AVLT, PVLT vs WVLT, and AVLT vs WVLT as Grouping Variables and the Measures of the VLT as Test Variables, with a Split File for Grade.

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Total score	Delayed recall	Recognition
P5								
PVLT-AVLT								
MW-U	115.000	129.500	143.000	161.000	178.000	127.000	129.500	133.000
Z	−2.451	−2.077	−1.712	−1.231	−0.799	−2.109	−2.075	−2.812
p	.014	.038	.087	.218	.436	.035	.038	.005*
PVLT-WVLT								
MW-U	135.000	148.500	149.000	166.000	152.500	140.500	135.500	169.000
Z	−1.408	−1.030	−1.016	−0.197	−0.916	−0.958	−1.400	−0.928
p	.159	.303	.310	.844	.359	.338	.162	.354
AVLT-WVLT								
MW-U	120.000	121.500	128.500	102.500	143.500	108.000	141.500	101.000
Z	−0.850	−0.792	−0.552	−1.93	−0.018	−0.979	−0.090	−1.915
p	.369	.428	.581	.233	.986	.328	.929	.056
P7								
PVLT-AVLT								
MW-U	28.500	33.000	113.500	122.500	109.500	56.500	83.500	79.500
Z	−4.265	−4.134	−1.567	−1.277	−1.681	−3.345	−2.523	−3.226
p	< .000*	< .000*	.117	.202	.093	.001*	.012	.001*
PVLT-WVLT								
MW-U	45.500	63.500	155.000	133.000	127.500	69.500	61.500	163.000
Z	−3.992	−3.448	−0.747	−1.398	−1.568	−3.237	−3.519	−0.935
p	< .001*	.001*	.445	.162	.117	.001*	< .001*	.350
AVLT-WVLT								
MW-U	122.500	167.000	152.000	177.000	152.500	157.000	151.000	102.500
Z	−1.721	−0.386	−0.839	−0.090	−0.820,	−0.674	−0.856	−2.713
p	.085	.699	.402	.929	.412	.500	.393	.007*
S1								
PVLT-AVLT								
MW-U	52.000	81.500	84.000	102.000	91.000	77.000	75.500	81.500
Z	−2.413	−1.046	−0.944	−0.134	−0.642	−1.228	−1.312	−1.562
p	.016	.296	.345	.893	.521	.219	.189	.118
PVLT-WVLT								
MW-U	67.000	53.500	61.500	94.000	56.500	55.000	74.000	104.500
Z	−1.688	−2.290	−1.939	−0.497	−2.218	−2.196	−1.373	−0.050
p	.091	.022	.052	.620	.027	.028	.170	.960
AVLT-WVLT								
MW-U	91.500	82.500	80.500	86.000	66.500	78.500	96.500	76.500
Z	−0.309	−0.723	−0.820	−0.565	−1.512,	−0.900	−0.070	−1.484
p	.758	.469	.412	.572	.131	.368	.944	.138
S2								
PVLT-AVLT								
MW-U	21.500	54.500	90.000	99.000	103.500	39.000	71.500	89.500
Z	−3.555	−2.243	−0.672	−0.270	−0.067	−2.709	−1.482	−1.132
p	< .000*	.025	.502	.788	.946	.007*	.128	.258
PVLT-WVLT								
MW-U	72.500	95.000	112.500	104.500	134.000	111.000	112.500	137.500
Z	−2.618	−1.858	−1.296	−0.325	−0.544	−1.304	−1.270	−0.756
p	.009*	.063	.195	.745	.587	.192	.204	.450

(*Continued*)

Table 5 (Continued).

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Total score	Delayed recall	Recognition
AVLT-WVLT								
MW-U	91.500	131.500	133.500	124.500	130.500	101.000	132.500	131.000
Z	−1.448	−0.302	−0.237	−0.558	−0.340	−1.072	−0.267	−0.477
p	.148	.763	.813	.577	.734	.284	.790	.634
S3								
PVLT-AVLT								
MW-U	38.500	75.000	73.500	65.500	60.000	54.000	70.000	77.000
Z	−1.748	−0.111	−0.196	−0.639	−0.956	−0.740	−0.389	.000
p	.080	.912	.845	.523	.339	.459	.697	1.000
PVLT-WVLT								
MW-U	36.000	34.500	24.000	33.500	35.000	35.500	31.500	33.000
Z	−0.232	−0.367	−1.368	−0.461	−0.329	−0.272	−0.642	−1.254
p	.817	.713	.171	.644	.742	.786	.521	.210
AVLT-WVLT								
MW-U	26.000	40.000	38.500	34.000	65.500	36.000	46.000	42.000
Z	−1.377	−0.681	−0.799	−1.107	−0.639	−1.007	−0.230	−1.414
p	.168	.496	.425	.268	.523	.314	.818	.157

Note. * = significant. MW-U = Mann Whitney U. For all significant findings, PVLT > AVLT, PVLT > WVLT, and WVLT > AVLT.

in grade S3 (wave 2) showed this effect also in wave 1 as the children in grade S2 at that time, but did not show these results in the second wave. Nevertheless, a small tendency towards differences for better PVLT performance as compared to AVLT performance was found for children in grade S3 on trial 1 (see Table 3), which could indicate a similar increase in the pictorial superiority effect. The remaining findings showed a few inconsistencies that we assume to be type-I errors.

Furthermore, the findings of the current study suggest that the development of the phonological loop and, more specifically, the (sub)vocal rehearsal system is the same in boys and girls. Girls and boys did not differ in their performance, regardless of presentation modality, grade, or verbal learning measure, except for delayed recall capacity in children below the age of 7 years, where girls outperformed boys. As described above, during delayed recall, the information needs to be actively recalled and searched in memory. Thus, girls in grade P2 might be better in that unguided search and retrieval in conditions with a time delay than boys, but in older children, this sex effect disappears. Thus, girls might indeed be faster than boys in their development of the retrieval of words after a prolonged period whereas boys catch up with girls above the age of 7 years. However, boys and girls do not seem to differ in their encoding and storage abilities, since no sex differences were found on recognition. These findings show that sex differences are more nuanced than most studies report with respect to VLT performance. This suggests that younger girls do indeed outperform boys on delayed recall, a finding that might be related to a better development of strategy use by girls (Cox & Waters, 1986), and this is line with studies that report no sex differences for the other measures of verbal learning (Alloway et al., 2006; Gathercole et al., 2004). However, based on these on these findings, we cannot rule out that sex differences are marginal or present but too subtle to be noticed, as has been suggested by other studies (Kramer et al., 2003; Van Den Burg &

Kingma, 1999). Indeed, more research in which the involvement of the phonological loop is actually manipulated is needed to confirm the findings of the current study.

The results the current study suggest that the development of the (sub)vocal rehearsal system in the phonological loop is a gradual process that develops around the age of 7 years and which is similar for boys and girls. Additionally, auditory information and textual information seemed to be processed in a similar manner, likely without labeling or recoding, leading to one processing stream, as compared to the dual coding of pictorial information. This leads to the overall conclusion that the beneficial effect of pictorial over auditory and textual information seems to exist if information is presented once (or at most twice) and only in children above the age of 7 years, and that this effect increases with age.

Original manuscript received 5 December 2014
Revised manuscript accepted 19 April 2015
First published online 19 June 2015

REFERENCES

Alloway, T. P., Gathercole, S. E., Kirkwood, H., & Elliot, J. (2009). The cognitive and behavioral characteristics of children with low working memory. *Child Development*, *80*, 606–621. doi:10.1111/cdev.2009.80.issue-2

Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, *77*(6), 1698–1716. doi:10.1111/cdev.2006.77.issue-6

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, *87*, 85–106. doi:10.1016/j.jecp.2003.10.002

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders*, *36*, 189–208. doi:10.1016/S0021-9924(03)00019-4

Baddeley, A. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. doi:10.1146/annurev-psych-120710-100422

Baddeley, A., & Hitch, G. J. (2000). Development of working memory: Should the Pascual-Leone and the Baddeley and Hitch models be merged? *Journal of Experimental Child Psychology*, *77*, 128–137. doi:10.1006/jecp.2000.2592

Bjorklund, D. F., & Douglas, R. N. (1997). The development of memory strategies. In N. Cowan & C. Hulme (Eds.), *The development of memory in childhood* (pp. 201–246). Hove, East Sussex: Psychology Press.

Cherry, K. E., Hawley, K. S., Jackson, E. M., Volaufova, J., Su, L. J., & Jazwinski, S. M. (2008). Pictorial superiority effects in oldest-old people. *Memory*, *16*, 728–741. doi:10.1080/09658210802215534

Constantinidou, F., & Baker, S. (2002). Stimulus modality and verbal learning performance in normal aging. *Brain and Language*, *82*, 296–311. doi:10.1016/S0093-934X(02)00018-4

Constantinidou, F., Danos, M. A., Nelson, D., & Baker, S. (2011). Effects of modality presentation on working memory in school-age children: Evidence for the pictorial superiority hypothesis. *Child Neuropsychology*, *17*, 173–196. doi:10.1080/09297049.2010.525503

Cowan, N. (1998). Visual and auditory working memory capacity. *Trends in Cognitive Sciences*, *2*, 77–78.

Cox, D., & Waters, H. S. (1986). Sex differences in the use of organization strategies: A developmental analysis. *Journal of Experimental Child Psychology*, *41*, 18–37. doi:10.1016/0022-0965(86)90048-2

Croy, C. D., & Novins, D. K. (2005). Methods for addressing missing data in psychiatric and developmental research. *Journal of the American Academy of Child & Adolescent Psychiatry*, *44*, 1230–1240. doi:10.1097/01.chi.0000181044.06337.6f

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177–190. doi:10.1037/0012-1649.40.2.177

Halliday, M. S., Hitch, G. J., Lennon, B., & Pettipher, C. (1990). Verbal short-term memory in children: The role of the articulator loop. *European Journal of Cognitive Psychology*, *2*, 23–38. doi:10.1080/09541449008406195

Henry, L. A., Turner, J. E., Smith, P. T., & Leather, C. (2000). Modality effects and the development of the word length effect in children. *Memory*, *8*, 1–17. doi:10.1080/096582100387678

Hitch, G. J., Halliday, M. S., Schaafstal, A. M., & Heffernan, T. M. (1991). Speech, 'inner speech', and the development of short-term memory: Effects of picture-labeling on recall. *Journal of Experimental Child Psychology*, *51*, 220–234. doi:10.1016/0022-0965(91)90033-O

Hitch, G. J., Woodin, M. E., & Baker, S. (1989). Visual and phonological components of working memory in children. *Memory & Cognition*, *17*, 175–185. doi:10.3758/BF03197067

Huizingh, E. (2002). *Inleiding SPSS 11 voor Windows* [Introduction SPSS for windows]. Schoonhoven: Academic Service.

Kemps, E., De Rammelaere, S., & Desmet, T. (2000). The development of working memory: Exploring the complementarity of two models. *Journal of Experimental Child Psychology*, *77*, 89–109. doi:10.1006/jecp.2000.2589

Kramer, J. H., Yaffe, K., Lengenfelder, J., & Delis, D. C. (2003). Age and gender interactions on verbal memory performance. *Journal of the International Neuropsychological Society*, *9*, 97–102. doi:10.1017/S1355617703910113

Lezak, M. D. (1995). *Neuropsychological assessment* (3rd ed.). New York, NY: Oxford University Press.

Lezak, M. D., Howieson, D. B., & Loring, D. W. (2004). *Neuropsychological assessment* (4th ed.). New York, NY: Oxford University Press.

Linschoten, J. (1963). *De La Court's Frequentietellingen van Nederlandse Woorden* [De la Courts' frequency counts of Dutch words]. Utrecht: University of Utrecht, Psychologisch Laboratorium.

Magimairaj, B. M., & Montgomery, J. W. (2012). Children's verbal working memory: Relative importance of storage, general processing speed, and domain-general controlled attention. *Acta Psychologica*, *140*, 196–207. doi:10.1016/j.actpsy.2012.05.004

Nevo, E., & Breznitz, Z. (2013). The development of working memory from kindergarten to first grade in children with different decoding skills. *Journal of Experimental Child Psychology*, *114*, 217–228. doi:10.1016/j.jecp.2012.09.004

Paivio, A. (1969). Mental imagery in associative learning and memory. *Psychological Review*, *76*, 241–263. doi:10.1037/h0027272

Palmer, S. (2000). Working memory: A developmental study of phonological recoding. *Memory*, *8*, 179–193. doi:10.1080/096582100387597

Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Tam, H., Jarrold, C., Baddeley, A. D., & Sabatos-DeVito, M. (2010). The development of memory maintenance: Children's use of phonological rehearsal and attentional refreshment in working memory tasks. *Journal of Experimental Child Psychology*, *107*, 306–324. doi:10.1016/j.jecp.2010.05.006

Thompson, V. A., & Paivio, A. (1994). Memory for pictures and sounds: Independence of auditory and visual codes. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, *48*, 380–398. doi:10.1037/1196-1961.48.3.380

Toglia, M. P., Hinman, P. J., Dayton, B. S., & Catalano, J. F. (1997). The blocked-random effect in pictures and words. *Perceptual and Motor Skills*, *84*, 976–978. doi:10.2466/pms.1997.84.3.976

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language*, *54*, 68–80. doi:10.1016/j.jml.2005.06.003

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology*, *71*, 1–26. doi:10.1016/j.cogpsych.2014.01.003

Vakil, E., Greenstein, Y., & Blachstein, H. (2010). Normative data for composite scores for children and adults derived from the Rey Auditory Verbal Learning Test. *The Clinical Neuropsychologist*, *24*, 662–677. doi:10.1080/13854040903493522

Van Den Burg, W., & Kingma, A. (1999). Performance of 225 Dutch school children on Rey's Auditory Verbal Learning Test (AVLT): Parallel test-retest reliabilities with an interval of 3 months and normative data. *Archives of Clinical Neuropsychology*, *14*, 545–559.

Van der Elst, W., Van Boxtel, M. P. J., Van Breukelen, G. J. P., & Jolles, J. (2005). Rey's Verbal Learning test: Normative data for 1,855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. *Journal of the International Neuropsychological Society*, *11*, 290–302. doi:10.1017/S1355617705050344

Van Loon-Vervoorn, W. A. (1989). *Voorstelbaarheidswaarden van Nederlandse woorden* [Imaebility ratings of Dutch words]. Lisse: Swets & Zeitlinger.

Whitehouse, A. J. O., Maybery, M. T., & Durkin, K. (2006). The development of the picture-superiority effect. *British Journal of Developmental Psychology*, *24*, 767–773. doi:10.1348/026151005X74153